Research Into Practice

Outcomes From a School-Randomized Controlled Trial of Steps to Respect: A Bullying Prevention Program

Eric C. Brown
University of Washington

Sabina Low
Wichita State University

Brian H. Smith
*Committee for Children
Seattle, WA*

Kevin P. Haggerty
University of Washington

Abstract. This study reports the outcomes of a randomized controlled trial of Steps to Respect: A Bullying Prevention Program conducted in 33 California elementary schools. Schools were matched on school demographic characteristics and assigned randomly to intervention or waitlisted control conditions. Outcome measures were obtained from (a) all school staff; (b) a randomly selected subset of third-, fourth-, and fifth-grade teachers in each school; and (c) all students in classrooms of selected teachers. Multilevel analyses indicated significant ($p <$

.05) positive effects of the program on a range of outcomes (e.g., improved student climate, lower levels of physical bullying perpetration, less school bullying-related problems). Results of this study support the program as an efficacious intervention for the prevention of bullying in schools.

Bullying is recognized as one of the most significant public health concerns facing children in the United States today and may be the most prevalent type of school violence (Porter, Batsche, Castillo, & Witte, 2006; Card & Hodges, 2008). It occurs along a continuum, with students assuming roles that include bully, victim, and bully-victim (Espelage & Horne, 2008). Bullying can result in negative psychological, emotional, and behavioral outcomes (Cook, Williams, Guerra, Kim, & Sadek, 2010). Victims, bullies, and bully-victims often report adverse psychological effects and poor school adjustment as a result of their involvement in bullying, which also might lead to subsequent victimization or perpetration (Nansel, Haynie, & Simons-Morton, 2003). For example, victims of bullying evidence more loneliness and depression, greater school avoidance, more suicidal ideation, and less self-esteem than their nonbullied peers (Hawker & Boulton, 2000; Kaltiala-Heino, Rimpelä, Rantanen, & Laippala, 2001; Kochenderfer & Ladd, 1996; Olweus, 1992; Rigby, 2001). Whereas victims tend to report more internalizing behaviors, bullies are more likely than their peers to exhibit externalizing behaviors, conduct problems, and delinquency (Haynie et al., 2001; Nansel et al., 2001). Lastly, bullying, in its many forms, is a serious problem that can harm students' school performance in the form of school avoidance, lower levels of academic achievement, and more conflictual relations with teachers and students (Glew, Fan, Katon, Rivara, & Kernic, 2005; Nansel et al., 2003).

Taken together, the prevalence and social-psychological costs of bullying warrant public health attention and efforts to alleviate the suffering involved. Given the high prevalence and strong relationship of bullying to adverse mental health outcomes, evidence-based school prevention programs are of great importance to school psychologists and other mental health professionals working with school-based youth.

Effectiveness of Bullying Interventions

Previous evaluations of school-based bullying preventive interventions have found mixed results, as reflected in four relatively recent research reviews. For example, Smith, Schneider, Smith, and Ananiadou (2004) synthesized evaluation studies of whole-school bullying prevention programs on bullying victimization and perpetration, and found that across most studies outcomes were negligible or negative. Moreover, the studies did not replicate the strong positive results found in the original test of the Olweus (1993) program in Norway (Smith et al., 2004). Vreeman and Carroll's (2007) review of bullying prevention evaluations (multidisciplinary or "whole-school" interventions, social skills groups, mentoring, and social worker support studies) found the majority of studies did not show positive effects, but interventions focused on the whole school were more effective than interventions delivered through classroom curricula or social skills training alone.

Merrell, Gueldner, Ross, and Isava (2008) conducted a meta-analysis of 16 school bullying preventive interventions that used quasi-experimental or experimental designs and found positive effects for roughly one third of the outcomes measured across the selected studies. However, outcomes ranged from actual bullying victimization and perpetration to correlated risk and protective factors (e.g., depression, self-esteem). Notably, 1 of the 16 studies found positive effects on bullying perpetration and 6 studies found positive effects on bullying victimization (student self-report). Finally, a recent meta-analysis by Farrington and Ttofi (2009) concluded that "overall, school-based antibullying programs are effective in reducing bullying and victimization"

(p. 6). Their meta-analysis found bullying decreased an average of 20% to 23% and victimization decreased an average of 17% to 20% in experimental versus control schools. However, they also noted that the programs examined were more effective in Europe and mostly did not use experimental designs, and cautioned against drawing any firm conclusions given the paucity of rigorously designed randomized controlled trials.

One possible explanation for divergent findings across these reviews is the methodological flaws found in many previous evaluations of bullying prevention interventions. A recent review by Ryan and Smith (2009) examined the scientific rigor and quality of 31 published bullying prevention evaluations. Deficiencies were found across all studies in the specification of intervention components, evaluation design (e.g., statistical power, unit of randomization), statistical analysis (e.g., multilevel vs. single level), program implementation, choice and measurement of outcomes (e.g., bullying behaviors, attitudes, and school climate), or selection of informants. Many studies were noted as failing to have an explicit theory of change guiding the evaluation or not matching the theory of change to the analytic strategy.

Steps to Respect: A Social-Ecological Approach to Bullying Prevention

Steps to Respect: A Bullying Prevention Program (STR) is a school-based prevention program that is aligned with the social-ecological model of bullying, which views youth behavior as shaped by multiple factors within nested contextual systems (Committee for Children, 2001). The program targets multiple areas of the school environment through intervention components directed at the school, peer, and individual levels (Swearer, Espelage, Vaillancourt, & Hymel, 2010). School-wide components are intended to foster a positive school climate and positive norms through teacher and staff training focused on the creation of effective disciplinary policies, improved monitoring of students, and instruction on how to effectively intervene with stu-

dents involved in bullying situations. Classroom curricula target the upper three elementary grades and are intended to promote socially responsible norms and behavior and increase social-emotional skills. Lessons help students recognize bullying, increase empathy for students that are bullied, build friendship skills to increase protective social connections, improve assertiveness and communication skills to help students deter and report bullying, and teach appropriate bystander responses to bullying.

The underlying theory of the STR program is that peer attitudes, norms, and behaviors play an important role in determining and maintaining rates of bullying behavior. Because bullying is a social process strongly influenced by the reactions and behaviors of peers (Atlas & Pepler, 1998), the program seeks to change attitudes about the acceptability of bullying through clearly labeling bullying behavior as unfair and wrong, increasing empathy for students who are bullied, and educating students about their responsibilities as bystanders to bullying. Figure 1 illustrates the social-ecological nature of the program's theory of change, with intervention components at the individual, peer, and school levels. As shown in this figure, the program is designed to reduce bullying in part through decreasing peer reinforcement of bullying behavior through increased positive bystander behaviors such as ignoring bullying, supporting students who are bullied, intervening to stop bullying incidents, and reporting bullying to school staff.

Previous evaluations of the STR program have demonstrated less acceptance of bullying and greater bystander responsibility and perceived adult responsiveness among intervention students than control students (Frey et al., 2005). Moreover, students participating in STR showed less increase in observed physical and relational forms of bullying than control students (Frey et al., 2005; Low, Frey, & Brockman, 2010), but reduced relational aggression was limited to those with supportive friends. The depth of the survey design for the Frey et al. (2005) study (e.g., recording observed playground behaviors using handheld computers) restricted the number of schools
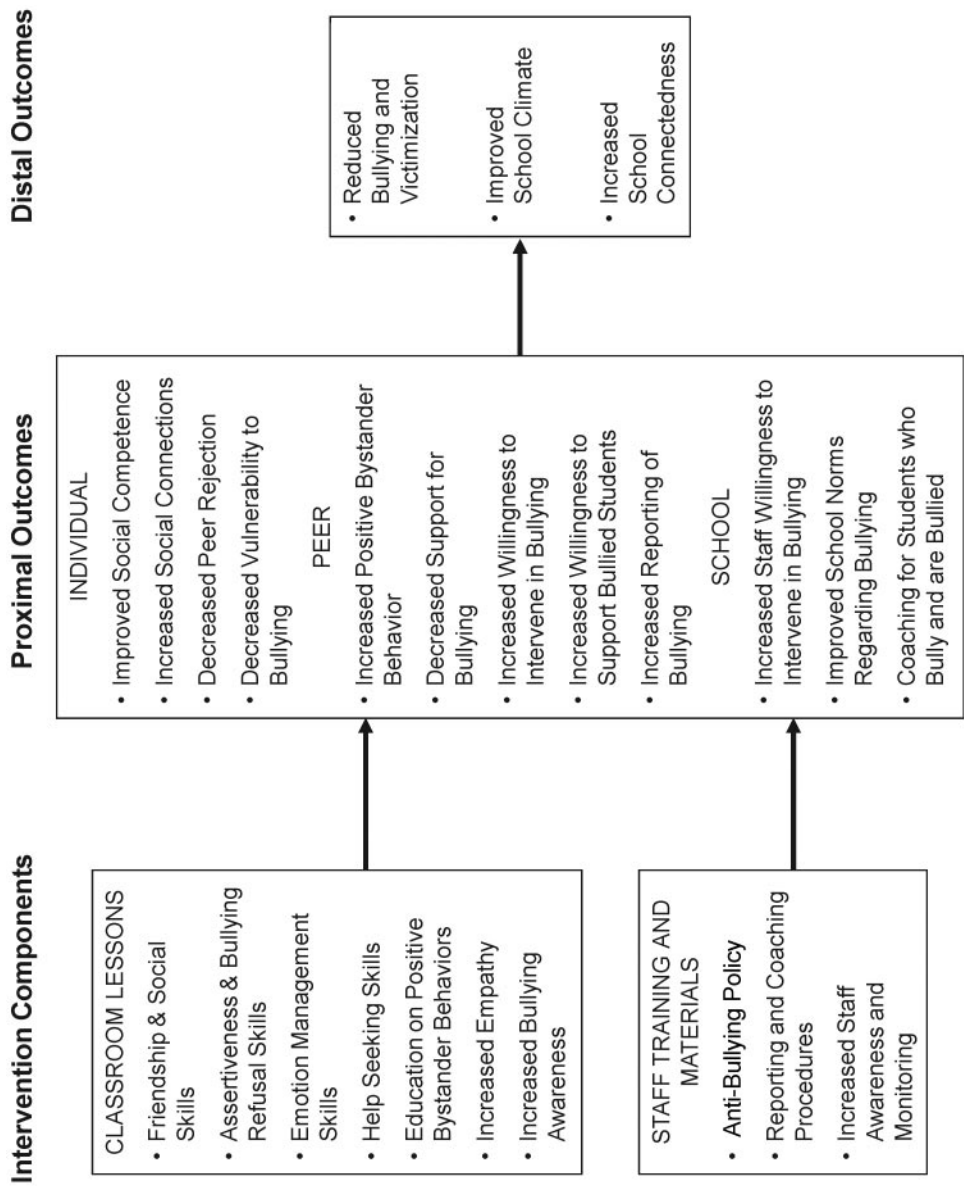
Figure 1. Theoretical framework for Steps to Respect: A School Bullying Prevention Program.

($N = 6$) that could be included in the trial. Nonetheless, significant effects were found for both observational and student self-report measures. These findings indicated promise for the efficacy of STR and prompted a call for a larger-scale school-randomized trial of the program.

Purpose

The purpose of the current study was to extend previous findings of STR program efficacy on reducing school bullying perpetration and victimization, and to assess the efficacy of the program to positively affect both proximal and distal bullying-related risk factors and outcomes, antibullying attitudes, social skills, bystander behaviors, and improved school climate and school connectedness. Our goal was to rigorously test the efficacy of the program using a sufficiently large number of schools and analysis of intervention effects at the school level, instead of student or classroom level, thus avoiding some the design limitations that have been noted in previous evaluations of school bullying prevention programs. Moreover, we sought to assess the efficacy of the program in light of the social-ecological nature of the program. Therefore, the following research questions guided the study and relied on data from multiple informants (i.e., students, teachers, and all school staff) to triangulate on program efficacy: (a) What are the effects of the STR program on the social-ecological context of the school? (b) What are the effects of the STR program on teacher perceptions of student behavior? (c) What are the effects of the STR program on student perceptions of school climate, staff/teachers, and connectedness with their school?

Methods

Participants

Schools. To select schools for participation, e-mails were sent to district contacts who met the following criteria: (a) had broad socioeconomic and racial/ethnic diversity, (b) had an established liaison with the Committee for Children, (c) expressed a strong need or desire for school bullying prevention, and (d) were not currently using a school bullying prevention program. To minimize expense, we limited schools to three geographic areas in north-central California. Day schools, alternative schools, and private and parochial schools were excluded from the eligible pool of elementary schools. Based on this initial effort, 45 out of 96 districts that expressed interest were selected for inclusion (i.e., 51 districts were not considered because they had a relatively small number of elementary schools, had no administrative contact, or were already using STR or another bullying-related program).

Research staff from Committee for Children visited the eligible elementary schools to introduce the study, explain the requirements, and answer any questions from principals or school staff. Principals from participating schools signed a statement that indicated that at least 80% of teachers in the schools agreed to participate in the program. Schools also agreed to receive program trainings for teachers and school staff, to implement the program as indicated in the trainings, and to be randomized into either intervention condition (i.e., beginning the program in the following school year) or waitlisted control condition (i.e., waiting 1 year before implementing the program). All participating elementary schools were provided with the STR program and trainings at no cost.

Based on power analyses using parameter estimates from Frey et al. (2005), and Jensen and Dietrich (2007), we targeted 34 schools for the trial. The 34 schools were matched into pairs within each geographic area using National Center for Education Statistics (n.d.) data on the characteristics of the school environment (e.g., total student enrollment, change in student enrollment from 2006 to 2007, number of teachers) and characteristics of the student population (e.g., percentage eligible for the free or reduced-price lunch program, ethnic/racial percentages, and percentage of students for whom English was not their primary language). Schools within each matched pair were assigned randomly to either

intervention or waitlisted control condition using a random number table.

Between the periods from random assignment to program implementation, two schools withdrew from the study (one because of turnover in leadership, and one because of building remodeling). One school was immediately replaced by another school from the original pool of eligible elementary schools, which was an adequate match on all criteria. However, a replacement was not found for the second school, leaving us with 33 schools for the trial. Twenty-five percent of the schools were from rural areas, 10% were from small towns, 50% were from suburban areas, and 15% were located in midsized cities. Schools had a mean of 40% of students receiving free or reduced-price lunch ($SD = 29\%$, range = 0 to 99%). The mean number of students per school was 479 ($SD = 177$, range = 77 to 749 students) and the mean number of teachers per school was 24.

School staff. School staff participants included all paid and volunteer staff from the 33 participating schools, including administrators, teachers, paraprofessionals, support staff, custodial and cafeteria personnel, bus drivers, and so on. At pretest, 1,307 individuals completed a survey (77% of the total population of school staff). At post-test, 1,296 individuals (76%) completed a survey. Respondents represented school administrators (2.8%), teachers (58%), paraprofessionals (10%), cafeteria staff (3.3%), school counselors/psychologists (1.4%), custodial staff (1.4%), bus drivers (0.7%), and other positions (7.7%). School staff participants were 90% female. Eighty-five percent of staff identified themselves as "Spanish, Latino, or Hispanic." Staff were 85% White, 2% African American, 4% Asian, and 9% Other race/ethnicity. School staff averaged 46 years of age and worked at their schools a median of 3–5 years.

Teachers. From each of the 33 participating schools, we randomly selected four third-, fourth-, or fifth-grade classrooms for data collection. Because the STR program was designed originally to be implemented to stu-

dents during the course of 3 consecutive years (although schools can implement the program for only 2 years, in Grades 5 and 6, if desired), our preference was to select third- and fourth-grade classrooms to allow for the potential of long-term follow-up; however, the limited number of third- and fourth-grade classrooms available to participate in the study required us to include fifth-grade classrooms in the study. Two schools had only three classrooms in this grade range and one school had only two classrooms in this grade range, leaving 128 classrooms: $n = 52$ third grade, $n = 62$ fourth grade, $n = 11$ fifth grade, $n = 2$ third- and fourth-grade split classrooms, and $n = 1$ fourth- and fifth-grade split classroom.

Students. All students in each of the selected classrooms were included in the target sample of 3,119 students for completion of the student survey. Students were given a study information sheet to be taken home and given to their parents at the beginning of the school year. Study information sheets informed parents about the study's goals, benefits, and possible risks of their child's participation in the study. Twenty-two parents declined participation in the study. An additional 173 students were determined to be ineligible because of language barriers or developmental disabilities, and an additional 14 students declined to participate in the study at either pretest or post-test administrations of the survey. Thus, the final analysis sample for pre–post outcome analyses consisted of 2,940 students (94% of the target population of students).

Among students in the intervention condition, approximately half (49%) were male, 52% were White, 7% were African American, 6% were Asian American, 35% were of other or mixed race. Forty-three percent of intervention students were identified as of Hispanic origin and the mean age of students in the intervention condition was $M = 8.9$ years ($SD = 0.84$, range = 7 to 11). Among control students, 52% were male, 53% were White, 6% were African American, 6% were Asian American, and 35% were of other or mixed race. Forty-one percent of control students were identified as of Hispanic origin and the

mean age of students in the control condition was $M = 8.9$ years ($SD = 0.81$, range = 7–11). All students were given a small gift worth approximately $5 each for their participation in the study. Procedures for passive consent for student participation in the study by students' parents were reviewed and approved by the University of Washington's Institutional Review Board.

Measures

School Environment Survey. Pretest and post-test data were collected from school staff during school staff training sessions (in intervention schools only) or during in-service meetings using the School Environment Survey (SES). The SES is a brief (10-min), anonymous, paper-and-pencil survey, which was adapted for the current study from the Colorado Trust's 3-year statewide Bullying Prevention Initiative (Csuti, 2008a). The SES was designed to parallel several of the measures collected from the student surveys to provide an alternative source of information on the social-ecological context of the school environment. The six SES outcome measures, number of items per scale, scale coefficient alphas (from the current sample), sample items, and anchors for response options are presented in Table 1.

School staff were asked about their perceptions of their school's climate regarding trust, willingness to help, and cooperation among students (Student Climate) and school staff (Staff Climate); willingness for students (Student Bullying Intervention) and staff (Staff Bullying Intervention) to intervene in observed bullying; perceptions of bullying-related problems among students in their school (School Bullying-Related Problems); antibullying policies and strategies in their school (School Antibullying Policies and Strategies); and background demographic information (age, gender, race/ethnicity, how many years they worked at the school, and their position at the school). School staff members were instructed to respond to survey items using the time frame "since the beginning of the school year." Scaled scores for outcome measures

were created as the mean of all nonmissing items on the scale and were considered missing if more than 67% of items in that scale were missing.

Teacher Assessment of Student Behavior (TASB). The TASB was a brief online survey of students' classroom behavior, scholastic aptitudes, and student demographic information developed by study investigators after a review of related measures. Teachers completed a separate online questionnaire for each student in their class. The five outcomes measured in the TASB, the number of items per scale, scale coefficient alphas (from the current sample), sample items, and anchors for response options are shown in Table 2.

TASB measures were identified to assess STR program efficacy in school and classroom competencies in interpersonal social skills (Social Competency), academic skills (Academic Competency), academic achievement (Academic Achievement); and teacher-observed physical (Physical Bullying Perpetration) and nonphysical (Nonphysical Bullying) acts of bullying. Teachers were instructed on the TASB to respond to survey items using the time frame "since the beginning of the school year." Scaled scores for TASB measures were created as the mean of all nonmissing items on the scale and were considered missing if more than 67% of items in that scale were missing.

Student survey. Similar to the SES, student data were collected using a revised version of the Colorado Trust's Bullying Prevention Initiative Student Survey (Csuti, 2008b). The 13 outcomes measured in the Student Survey, the number of items per scale, sample items, scale coefficient alphas (from the current sample), and anchors for response options are presented in Table 3.

In addition to student demographic information (age, gender, and race/ethnicity), students were asked about their perceptions of their school's climate regarding trust, willingness to help, and cooperation among fellow students (Student Climate) and school staff/teachers (Staff Climate); and about their con-

Table 1
Characteristics of School Environment Survey Outcome Measures

Outcome	Number of Items	Coefficient α	Sample Item	Response Options Range and Anchors
School Antibullying Policies and Strategies	8	.93	How much is your school doing to demonstrate administrator commitment and leadership to address bullies, bullied, and bystanders?	1 = *Not at all* to 4 = *A lot*
Student Bullying Intervention	5	.92	Students in this school would help out if a student is making fun of or teasing another student who is obviously weaker.	1 = *Never* to 7 = *Every time*
Staff Bullying Intervention	5	.95	Teachers and staff in this school would help out if a student or group of students is trying to pick a fight with a weaker student.	1 = *Never* to 7 = *Every time*
Student Climate	4	.82	Students in this school are willing to help other students.	1 = *Strongly disagree* to 4 = *Strongly agree*
Staff Climate	7	.91	Staff in this school can be trusted.	1 = *Strongly disagree* to 4 = *Strongly agree*
School Bullying-Related Problems	7	.82	How big of a problem in your school is students spreading rumors or lies about students they are mad at or don't like?	1 = *Not a problem* to 4 = *A huge problem*

Table 2
Characteristics of Teacher Assessment of Student Behavior Outcome Measures

Outcome	Number of Items	Coefficient α	Sample Item	Response Options Range and Anchors
Social Competency	5	.82	Gets along with classmates.	1 = *Never* to 5 = *Always*
Academic Competency	4	.86	Needs help to stay on task.	1 = *Never* to 5 = *Always*
Academic Achievement	3	.95	Academically, how would you rate this student in terms of reading?	1 = *Needs much improvement* to 5 = *Above average*
Physical Bullying Perpetration	4	.91	Pushed, shoved, or tripped a weaker student.	1 = *Never* to 5 = *Always*
Nonphysical Bullying	4	.80	Spread rumors about another student.	1 = *Never* to 5 = *Always*

nectedness or bonding with their school (School Connectedness). Additional Student Survey measures asked students about their perceptions of general student support and empathy (Student Support), their attitudes regarding normative beliefs against acts of bullying perpetration (Attitudes Against Bullying) and bullying intervention (Attitudes Toward Bullying Intervention), student and teacher/staff willingness to intervene in observed acts of bullying (Student Bullying Intervention and Teacher/Staff Bullying Intervention, respectively), and whether teachers and staff were doing the "right things" to prevent bullying (Teacher/Staff Bullying Prevention).

Other, more distal, outcomes assessed by the Student Survey included measures of appropriate (as defined by the STR program) bullying bystander behaviors (Positive Bystander Behavior), perceptions of bullying as a problem in their school (School Bullying-Related Problems), and observed instances of actual bullying perpetration (Bullying Perpetration) and victimization (Bullying Victimization) among students. Four scales—Student Attitudes Against Bullying, Students Attitudes Toward Bullying Intervention, Teacher/Staff Bullying Prevention, and Bullying Perpetration—exhibited a high degree of nonnormality. Given the response options for these scales, Student Attitudes Against Bullying and Students Attitudes Toward Bullying Intervention were operationalized as count variables; that is, mean values for Student Attitudes Against Bullying indicated average number of attitudes considered to be *Little wrong*, *Very wrong*, or *Extremely wrong*, and could range from 0 to 7 for the 7-item scale. Mean values for Student Attitudes Toward Bullying Intervention indicated average number of attitudes considered to be *Little okay* or *Very okay*, and could range from 0 to 4 for the 4-item scale. Teacher/Staff Bullying Prevention was analyzed as an ordered categorical outcome (Raudenbush & Bryk, 2002), and Bullying Perpetration was dichotomized into 0 = *Never* and 1 = *Sometimes* to *Always*. Students were instructed to respond to survey items using the time frame "since the beginning of the school year." Scaled scores for student survey out-

Table 3
Characteristics of Student Survey Outcome Measures

Outcome	Number of Items	Coefficient α	Sample Item	Response Options Range and Anchors
Student Support	6	.78	Students my age are there for me whenever I need help.	1 = *Not true at all to 4* = *Completely true*
Attitudes Against Bullying	7	.87	How wrong or okay is it when students push, shove, or pick fights with weaker students?	1 = *Extremely wrong to* 5 = *Very okay*
Attitudes Toward Bullying Intervention	4	.79	How wrong or okay is it when students defend others who are being shoved around by strong students?	1 = *Extremely wrong to* 5 = *Very okay*
Teacher/Staff Bullying Prevention	1	na	Teachers and staff at my school are doing the right things to prevent bullying.	1 = *Strongly disagree to* 4 = *Strongly agree*
Student Bullying Intervention	4	.76	Students in your school would help out if a student or group of students pushed, shoved or tried to pick a fight with a weaker student.	1 = *Never to 4* = *Always*
Teacher/Staff Bullying Intervention	4	.85	Teachers and staff in your school would help out if a student spread rumors or lies about another student behind their back.	1 = *Never to 4* = *Always*
Positive Bystander Behavior	5	.69	I tried to defend students who always get pushed or shoved around.	1 = *Never to 5* = *A lot*
School Bullying-Related Problems	7	.87	How big of a problem in your school is students picking fights with other students?	1 = *No problem at all to* 4 = *A huge problem*
Bullying Perpetration	7	.87	I teased or said mean things to certain students.	1 = *Never to 5* = *Always*
Bullying Victimization	4	.75	A particular student or group of students spread rumors or told lies about me.	1 = *Never to 5* = *Always*
Student Climate	4	.68	Students in my school can be trusted.	1 = *Strongly disagree to* 4 = *Strongly agree*
School Connectedness	5	.76	This is a pretty close-knit school where everyone looks out for each other.	1 = *Strongly disagree to* 4 = *Strongly agree*
Staff Climate	7	.82	Teachers and staff in my school usually get along with students.	1 = *Strongly disagree to* 4 = *Strongly agree*

Note. na = Not applicable.

comes were constructed as the mean of all nonmissing items. Scales missing more than 67% of the items were set to missing.

Procedure

Teachers from each classroom actively consented to their participation in the study. Participating teachers in intervention and control schools were reimbursed $5 per student for their time in completing the survey at pretest and post-test, with an additional bonus of $25 if all assessments were completed within 2 weeks of the initial announcement of the survey. Participating teachers in intervention schools were reimbursed an additional $75 if all 11 of the online Program Implementation Logs were completed by the end of the study.

Schools in this study received the standard, fully manualized, Steps to Respect program that is routinely offered to schools in regard to the classroom lessons, staff training, and support materials (see Web site for more details, http://www.cfchildren.org).

Staff training. Committee for Children trainers provided an on-site, one-day training for all participating teachers and staff. This training is typically offered to schools, as well as provided in the program materials for schools that wish to train their own staff. As part of this training, all staff in the school received a 3-hr overview of program goals and key features of program content (e.g., a definition of bullying, a model for responding to bullying reports). Teachers, counselors, and administrators received an additional 1.5-hr training in how to coach students involved in bullying. Third- through sixth-grade teachers also received a 2-hr overview of classroom materials and lesson-specific instructional strategies.

Classroom curriculum. Eleven semiscripted skills lessons focusing on social-emotional skills for positive peer relations; emotion management; and recognizing, refusing, and reporting of bullying behavior were delivered by teachers. Lesson topics included joining groups, distinguishing reporting from tat-

tling, and being a responsible bystander. Instructional strategies included direct instruction, large- and small-group discussions, skills practice, and games. Each of the weekly lessons, totaling about 1 hr, was taught over 2–3 days. There were no make-up sessions for students that missed a lesson. Upon completion of skill lessons, teachers implemented a grade-appropriate literature unit, based on existing children's books, which provided further opportunities to explore bullying-related themes.

Parent engagement. A scripted informational overview for parents was sent home with students. Take-home letters for parents, provided throughout the classroom curriculum, outlined key concepts and skills and described activities to support their use at home. Administrators informed parents about the program and the school's antibullying policy and procedures.

Implementation sequence. The STR program was implemented in several phases during the trial. First, in the fall of 2008, school bullying prevention teams met and collaborated with program consultants to develop the infrastructure to implement and/or sustain school prevention efforts (e.g., handling of reports and coaching for students involved in bullying). Second, in November of 2008, school personnel were trained in the STR program. Finally, classroom lessons were implemented in third- through six-grade from December 2008 through May 2009.

Implementation Fidelity

Teachers completed an online Program Implementation Log, which was to be completed at the end of every week in which a classroom lesson was supposed to be completed. Teachers' self-reported responses to the third- through fifth-grade ratings of schoolwide implementation using a 4-item scale (1 = *Poor*, 4 = *Excellent*) indicated that by the end of the school year program policies and procedures were well implemented ($M = 3.25$, $SD = 0.44$). Eighty-three percent of all teachers in intervention schools reported

teaching at least 80% of the lessons and 91% of teachers reported teaching at least 60% of the lessons. Teachers reported teaching 99.2% of all classroom skill lessons. Approximately 75% of students were exposed to at least 95% of all lessons, and an additional 22% of students were exposed to between 75% and 94% of the lessons. Overall, program engagement, identified as the average response across all weekly reported lessons to the question, "To what extent were your students engaged by this lesson? " (1 = *Not at all*, 4 = *A lot*), was high (*M* = 3.67, *SD* = 0.54). Across all lessons, 18% of teachers reported omitting one or more elements from a lesson. Despite this, a high percentage of teachers (92%) reported completing all objectives.

Statistical Analyses

Analyses consisted of mixed-model analysis of covariance, implemented as a hierarchical linear model (Raudenbush & Bryk, 2002) for continuous outcomes or as a hierarchical generalized linear model (Raudenbush, Bryk, Cheong, & Congdon, 2004) for binary, count, and ordered categorical outcomes. We modeled outcome measures from the TASB and Student Survey as three-level models, with students nested in classrooms, and classrooms nested within schools. The model controlled statistically for student baseline characteristics: age, gender, race/ethnicity (Black vs. White, Other race vs. White, and Hispanic vs. Non-Hispanic), grade (Grade 3 vs. Grade 4 and Grade 4 vs. Grade 5), and the pretest measure of the outcome. The number of students in the classroom was included as a control variable at Level 2. Two dummy variables representing the three geographic areas were included as control variables at Level 3. An additional dummy variable, coded 1 for intervention schools and 0 for waitlisted control schools, was included at Level 3. The intervention effect was estimated as the mean difference in adjusted school-level means (or proportions) between intervention and control schools tested against the average variation among the intervention-condition-specific adjusted school-level means (or proportions), expressed as robust standard errors, with degrees of freedom equal to the number of schools minus the number of school-level covariates and intervention effect, minus one (i.e., *df* = 29). Models for nonnormal outcomes incorporated the appropriate link function for the dependent variables (Raudenbush & Bryk, 2002). The models included random effects for intercept parameters across classrooms and schools. Intraclass correlation coefficients (ICCs) indicated the proportion of total variability in an outcome that could be attributed to each level.

Because the SES was an anonymous survey, it was not possible to link pretest and posttest surveys by individual staff persons. Thus, analysis of SES outcome data necessitated a two-level modeling strategy (see Livert, Rindskopf, Saxe, & Stirratt, 2001, for an example), with school staff members (Level 1) nested within schools (Level 2). Level 1 control variables included staff members' gender, race/ethnicity, age, duration of employment at the school, and whether they held an administrative or nonacademic position (relative to a teaching position, which was the reference category). Level 2 control variables included dummy variables for geographic area. In this model, an additional Level 1 covariate, Time, was included to denote whether the dependent outcome variable was measured at pretest (coded 0) or post-test (coded 1). The intervention effect in this model was identified as the regression of the Level 1 slope for Time on the Level 2 intervention status variable. In all models, effect sizes were calculated as the difference in post-test adjusted means (i.e., the intervention effect), divided by the square root of the sum variances across all respective levels (student, classroom, and school) of conditional hierarchical linear models (i.e., pooled total standard deviation). Using Cohen's (1988) standards, effect sizes less than 0.3 might be considered "small" effects, effect sizes around 0.5 might be considered "medium" effects, and effect sizes larger than 0.8 might be considered "large" effects. For dichotomous or categorical outcomes, effect sizes were calculated as adjusted odds ratios. All analyses were performed using the hierar-

chical linear models for Windows software program, version 6.06 (Raudenbush et al., 2004).

Baseline equivalency. Results of conditional models examining equivalency in baseline levels of outcome measures between intervention and control schools indicated that 1 of the 24 assessed outcomes across the three surveys demonstrated significantly different mean pretest levels. Students from intervention schools reported higher pretest levels of student bullying intervention than did students in control schools, $t(29) = 2.51, p < .05$.

Missing data. A statistical comparison of participants that had missing scale data versus those who did not have missing data on outcome scales indicated that rates of missingness did not differ by participants' gender, race/ethnic group, age, grade, or intervention status ($p > .05$). However, rates of missing scale data were significantly different by geographic area, with one area demonstrating lower rates of missing data than the other two geographic areas, $\chi^2(1, N = 3,048) = 4.81, p < .05$. Except for items that asked teachers and students about students' use of electronic devices to perpetrate bullying behaviors, overall rates of missing data were low (less than 10% across scales in any given survey). Nonetheless, missing data were accounted for using multiple imputation analyses (Graham, 2009) via NORM version 2.03 (Schafer, 1997). To preserve the unique variance–covariance structures of the data by intervention status, we conducted separate imputation analyses for experimental and control schools. Forty imputed data sets were created for each survey by intervention status group. Imputed data sets were combined subsequently to include both intervention and control groups for analysis of outcome measures.

Results

Social-Ecological Context of the School

Model-implied pretest and post-test means and standard deviations (by interven-tion condition), and tests of intervention effects for SES measures from the multilevel analyses, are presented in Table 4. Significant intervention effects were present for five of the six examined SES outcome measures, with results indicating greater increases in school antibullying policies and strategies, $t(29) = 3.33, p < .01$; student climate, $t(29) = 3.25, p < .01$; and staff climate, $t(29) = 2.91, p < .01$; less decrease in student bullying intervention, $t(29) = 3.42, p < .01$; and a larger decrease in school bullying-related problems, $t(29) = -2.91, p < .01$, for intervention schools relative to control schools. The average effect size across these five outcomes was 0.296 (range = 0.212 for staff climate to 0.382 for antibullying policies and strategies). No intervention effect was found for staff bullying intervention.

Among the pretest covariate effects, school staff who held administrative positions in schools reported higher levels of student climate, student bullying intervention, and school antibullying policies and strategies than did teachers, $t(2,588) = 2.00, 2.58,$ and $7.48,$ p values $< .05$, respectively. Teachers reported higher levels of school bullying-related problems than did either administrative personnel, $t(2,588) = -3.67, p < .001$, or school staff from nonacademically related positions, $t(2,588) = -2.20, p < .05$. Teachers also reported higher levels of staff bullying intervention, $t(2,588) = 2.06, p < .05$, and lower levels of school antibullying policies and strategies than did nonacademic staff, $t(2,588) = -5.15, p < .001$, respectively. Older staff reported greater staff climate and school antibullying policies and strategies, $t(2,588) = 2.49$ and $4.17, p$ values $< .05$, respectively; and less school bullying-related problems, $t(2,588) = -4.98, p < .001$, than younger staff. Staff members' length of employment was related positively to higher levels of student bullying intervention, $t(2,588) = 2.19, p < .05$; and negatively to school antibullying policies and strategies, $t(2,588) = -3.29, p < .01$. Conditional ICCs (i.e., including staff and school characteristics as covariates) corresponding to between-school variation averaged .096 and ranged

Table 4
Pretest and Post-Test Adjusted Means, Standard Deviations, and Tests of Intervention Effects for School Environment Survey Measures

| | Fall Pretest | | | | Spring Post Test | | | | | |
| | Intervention | | Control | | Intervention | | Control | | | |
Outcome	M	SD	M	SD	M	SD	M	SD	t df = 29	Effect Size
School Antibullying Policies and Strategies	2.84	0.764	2.88	0.73	3.18	0.65	2.94	0.76	3.33*	0.38
Student Bullying Intervention	4.47	1.05	4.67	1.07	3.88	1.30	3.75	1.26	3.42*	0.28
Staff Bullying Intervention	5.99	0.91	6.03	0.91	6.17	1.12	6.21	0.94	−0.34	na
Student Climate	3.07	0.48	3.13	0.46	3.16	0.47	3.10	0.45	3.25*	0.21
Staff Climate	3.45	0.46	3.47	0.47	3.51	0.47	3.41	0.49	2.91*	0.26
School Bullying-Related Problems	3.55	1.05	3.31	0.97	1.83	0.54	1.90	0.58	−2.91*	−0.35

Note. na = Not applicable. Effect size calculated as difference in post-test adjusted means divided by pooled total standard deviation from conditional model.
*$p < .01$.

from 5% of variance in the student bullying intervention to 16% of variance in antibullying policies and strategies being attributable to participants' schools.

Teacher Perception of Student Behavior

Model-implied pretest and post-test means and standard deviations and tests of intervention effects from the multilevel analyses are presented in Table 5 by intervention condition. Two of the five assessed teacher outcomes demonstrated significant intervention effects in the conditional HLMs. Whereas teachers from control schools reported declines in mean levels of social competency from pretest to post-test, teachers from intervention schools reported little change in this outcome, $t(29) = 2.16$, $p < .05$. In addition, although the prevalence of physical bullying perpetration increased during the school year in both control and intervention schools, the increase was smaller in intervention schools, $t(29) = -3.12$, $p < .01$. Effect sizes associated with these intervention effects were 0.131 for the standardized difference in Social Com-

petency means and adjusted odds ratio = 0.609, indicating a reduction of 31% in the likelihood of Physical Bullying Perpetration in intervention schools relative to control schools. No significant differences were found between intervention and control schools for nonphysical bullying perpetration, academic competency, or academic achievement.

Pretest covariate effects indicated that males were reported to be more likely to engage in physical bullying than females, $t(2,789) = 5.69$, $p < .001$; and to have lower levels of social competency and academic competency than females, $t(2,789) = -5.26$ and $t(2789) = -6.28$, p values $< .01$, respectively. African American students were reported to be more likely to engage in physical bullying and to have lower levels of social competency, academic competency, and academic achievement than White students, $t(2,789) = 2.56$, -2.87, and -2.79, p values $< .05$, respectively. Hispanic students also were reported to have lower levels of academic achievement than non-Hispanic students, $t(2,789) = -4.47$, $p < .001$. Older

Table 5
Pretest and Post-Test Adjusted Means, Standard Deviations, and Tests of Intervention Effects for Teacher Assessment of Student Behavior Measures

| | Fall Pretest | | | | Spring Post Test | | | | | |
| | Intervention | | Control | | Intervention | | Control | | | |
Outcome	M	SD	M	SD	M	SD	M	SD	t df = 29	Effect Size
Social Competency	4.04	0.67	4.04	0.621	4.05	0.72	3.98	0.67	2.16*	0.13
Academic Competency	4.03	0.92	4.11	0.885	4.02	0.91	4.01	0.88	0.47	na
Academic Achievement	3.03	1.26	3.08	1.25	3.18	1.25	3.19	1.26	−0.15	na
Physical Bullying Perpetration[a]	0.21	0.41	0.17	0.378	0.23	0.42	0.286	0.45	−3.12**	AOR = 0.61
Nonphysical Bullying Perpetration[a]	0.42	0.49	0.40	0.489	0.49	0.50	0.517	0.50	−1.36	na

Note. AOR = adjusted odds ratio; na = not applicable. Unless otherwise noted, effect size calculated as difference in post-test adjusted means divided by pooled total standard deviation from conditional model.
[a]Dichotomized into *Never* versus *Some* to *Always* (i.e., mean values indicate proportions of students with at least one bullying behavior).
*$p < .05$. **$p < .01$.

students were reported to be more likely to engage in physical and nonphysical bullying, $t(2,789) = 2.51$ and 3.31, p values $< .01$, respectively; and to have lower levels social competency and academic achievement than younger students, $t(2,789) = -3.42$ and -3.48, p values $< .01$, respectively.

Finally, class size was associated positively with levels of social competency, $t(2,789) = 1.98$, $p < .05$, and negatively with the prevalence of nonphysical bullying, $t(2,789) = -3.70$, $p < .01$. Conditional ICCs for between-classroom variation averaged .095 and ranged from 8% to 11% of variation among classrooms for academic achievement and social competency measures, respectively. Conditional ICCs variation among schools averaged .042 across the five outcomes, ranging from 4% to 5% for academic achievement and social competency, respectively.

Student Perceptions

Model-implied pretest and post-test means and standard deviations for intervention and control schools, and tests from the multilevel analysis of covariance model assessing intervention effects, are presented in Table 6. Significant intervention effects were found for 5 of the 13 student outcomes across the range of proximal and distal outcomes. For example, whereas students from intervention schools reported higher mean levels of student climate at post-test than at pretest, students from control schools reported lower post-test levels of this outcome than at pretest, $t(29) = 2.39$, $p < .05$. Students from intervention schools reported significantly less of a decline in teacher/staff bullying prevention during the school year, $t(29) = -2.22$, $p < .05$; and greater increases in student bullying intervention, teacher/staff bullying intervention, and positive bystander behavior, than did students from control schools, $t(29) = 2.35$, 2.54, and 2.62, p values $< .05$, respectively.

No significant differences between intervention and control schools were found for Student Support, Student Attitudes Against Bullying, Student Attitudes Toward Bullying Intervention, School Bullying-Related Problems, Bullying Perpetration, Bullying Victimization, School Connectedness, and Staff Cli-

Table 6
Pretest and Post-Test Adjusted Means, Standard Deviations, and Tests of Intervention Effects for Student Survey Measures

| | Fall Pretest | | | | Spring Post Test | | | | | |
| | Intervention | | Control | | Intervention | | Control | | | |
Outcome	M	SD	M	SD	M	SD	M	SD	t df = 29	Effect Size
Student Support	2.92	0.69	2.91	0.68	2.81	0.66	2.76	0.70	1.22	na
Student Attitudes Against Bullying[a]	6.18	1.68	6.19	1.73	5.64	2.13	5.55	2.23	−0.87	na
Student Attitudes Toward Bullying Intervention[a]	2.32	0.99	2.28	0.97	2.99	1.42	3.00	1.38	−0.21	na
Teacher/Staff Bullying Prevention[b]	3.28	0.92	3.26	0.94	3.23	0.90	3.10	0.99	−2.22*	AOR = 1.27
Student Bullying Intervention	2.86	0.83	2.81	0.83	2.94	0.77	2.83	0.76	2.35*	0.12
Teacher/Staff Bullying Intervention	2.86	0.98	2.82	0.97	3.05	0.89	2.94	0.91	2.54*	0.13
Positive Bystander Behavior	2.38	0.88	2.34	0.91	2.72	1.09	2.62	1.07	2.62*	0.14
School Bullying-Related Problems	2.68	0.875	2.73	0.88	2.60	0.83	2.68	0.82	−1.27	na
Bullying Perpetration[c]	0.50	0.500	0.468	0.50	0.584	0.49	0.60	0.49	−1.06	na
Bullying Victimization	2.14	1.04	2.10	1.04	2.11	1.03	2.18	1.06	−1.15	na
Student Climate	2.57	0.592	2.57	0.58	2.60	0.52	2.51	0.55	2.39*	0.19
School Connectedness	2.99	0.653	3.02	0.60	2.89	0.67	2.86	0.68	0.77	na
Staff Climate	3.15	0.525	3.18	0.50	3.05	0.58	3.08	0.57	−0.89	na

Note. AOR = adjusted odds ratio; na = not applicable. Unless otherwise noted, effect size calculated as difference in post-test adjusted means divided by pooled total standard deviation from conditional model.
[a]Operationalized as a count variable (i.e., mean values for Student Attitudes Against Bullying indicate average number of attitudes considered to be *Little wrong, Very wrong,* or *Extremely wrong;* mean values for Student Attitudes Toward Bullying Intervention indicate average number of attitudes considered to be *Little okay* or *Very okay*).
[b]Operationalized as ordered categorical variable. [c]Dichotomized into *Never* versus *Some* to *Always* (i.e., mean values indicate proportions of students with at least one bullying behavior).
*$p < .05$.

mate. Effect sizes associated with standardized differences in adjusted means between intervention and control schools for significant outcomes ranged from 0.115 for Student Bullying Intervention to 0.187 for Student Climate. For the ordinal teacher/staff bullying prevention outcome, the adjusted odds ratio was 1.27, indicating that the incremental likelihood of being more in agreement that teachers and staff were doing the right things to prevent bullying was 27% greater for students from intervention schools than students from control schools.

Pretest covariate effects indicated that male students reported lower levels of school connectedness, student attitudes against bullying, teacher/staff bullying intervention, and positive bystander behavior than female students, $t(2,836) = -2.46, -4.25, -3.79,$ and $-3.10, p$ values $< .05$, respectively. Male students also reported more bullying perpetration that did female students, $t(2,684) = 3.13,$ $p < .05$. African American students reported lower levels of student climate, $t(278) =$ $-2.49, p < .05,$ and higher levels of school bullying-related problems and bullying perpe-

tration than White students, $t(3,035) = 3.08$ and 2.61, p values < .01, respectively. Similarly, students from other (non-African American) minority groups also reported higher levels of school bullying-related problems, $t(2,201) = 3.70$, $p < .001$, and lower levels of school connectedness and teacher/staff bullying prevention than White students, $t(712) = -2.47$ and -3.29, p values < .05, respectively. Hispanic students reported lower levels of positive bystander behavior, teacher/staff bullying intervention, and bullying victimization than non-Hispanics, $t(2,089) = -2.97$, -2.29, and -1.96, p values < .05, respectively, and higher levels of bullying perpetration than non-Hispanic students, $t(1,468) = 2.35$, $p < .05$.

Younger students reported higher levels of student climate, staff climate, and school connectedness than did older students, $t(1,644) = -3.09$, -2.20, and -1.98, p values < .05, respectively. Older students reported higher levels of student attitudes toward bullying intervention, bullying victimization, and bullying perpetration that did younger students, $t(2,684) = 3.12$, 2.93, and 3.12, p values < .05, respectively. Finally, the number of students in the classroom was associated positively with students' perceptions of school bullying-related problems and student attitudes against bullying, $t(126) = 2.40$ and 2.06, p values < .05, respectively. Conditional ICCs for variation among classrooms averaged .022 and ranged from 1% for student bullying intervention to 4% for teacher/staff climate. Conditional ICCs for variation among schools averaged .019 and ranged from 2% for teacher/staff bullying prevention to 2% for bullying-related problems.

Discussion

This article reports results of a school-randomized controlled trial of Steps to Respect: A Bullying Prevention Program. Our study employed a rigorous experimental design that randomized 33 elementary schools to intervention or waitlisted control condition and used appropriate multilevel analyses that accounted for student-, classroom-, and school-level effects. The study design, statistical analyses, and survey battery were consistent with recommendations in the field for assessing the effects of interventions that target the whole-school environment (Bloom, Bos, & Lee, 1999; Raudenbush & Willms, 1995), and for evaluating school bullying prevention programs, in particular (Ryan & Smith, 2009).

Baldry and Farrington (2007) recommended that school bullying prevention programs be based on strong program theory. Thus, our evaluation was consistent with a social-ecological approach to school bullying prevention, which suggests that prevention effectiveness is maximized when intervention occurs at multiple levels, including individual students, their peers, and the larger school environment. Accordingly, we obtained information from multiple sources within each school (i.e., from students, teachers, and school staff), across domains specified in the theory, and subsequently analyzed our data to reflect the social-ecological nesting.

Results of this study demonstrated significant intervention effects for the prevention of school bullying on 50% of all outcomes examined across the three sources of data. Moreover, intervention effects were found for both proximal and distal outcomes. Following Cohen's (1988) guidelines for interpreting the magnitude of standardized between-group differences, most observed intervention effect sizes were relatively small (i.e., less than 0.3); however, effects of this magnitude are not unexpected given the short duration (i.e., 1 year) of the study. Long-term follow-up of students in a multiyear longitudinal study, with full dosage and exposure of students to the intervention, might show larger effect sizes (see, for example Frey, Hirschstein, Edstrom, & Snell, 2009). Moreover, increases in normative classroom aggression and related bullying behaviors have been noted during the school year (Frey et al., 2005) and it is important to contextualize intervention effects in light of this normative escalation. Nonetheless, the small effect sizes reported in this study may be seen as a limitation to the practical significance of study findings.

Additional effects of model covariates indicated differences in mean levels of bullying-related behaviors across gender, racial/ethnic, and age groups. Although not central to aims of the current study, these differences are interesting and support existing literature on the characteristics of students that are associated with bullying. For example, higher rates of bullying perpetration and bullying-related problems among African American, Hispanic, and other minority students are consistent with Graham and Juvonen's (2002) findings of minority ethnic group differences on aggressive behavior. Higher rates of bullying perpetration found for male students are consistent with Espelage et al. (2004), Nansel et al. (2001), and Seals and Young (2003). Future studies of school bullying prevention program effectiveness should take these differences into consideration and statistically control for them in outcome analyses.

In addition to intervention and covariate effects, our results include estimates of intraclass correlations of classroom and school variability to help researchers plan future school-randomized trials of bullying prevention programs. Observed intraclass correlation coefficients indicated that appreciable levels of variation in staff and teacher reported outcomes exist at both classroom and school levels. This suggests that both the classroom and the entire school building are viable, perhaps even necessary, environments for preventive intervention.

Consistent with the STR program's theory of change, we found significant intervention effects on outcome measures from multiple domains of the school environment, including school, peer, and individual student domains. For example, among proximal outcomes, effects on bullying prevention efficacy in the school domain were found in School Antibullying Policies and Strategies (from the SES), and in Teacher/Staff Bullying Prevention and Teacher/Staff Bullying Intervention measures (Student Survey); effects among peers were found in Positive Bystander Behaviors (Student Survey) and Student Bullying Intervention (SES and Student Surveys); and effects among students were found in Social

Competency (TASB). Among distal outcomes, intervention efficacy in the school domain was found in School Bullying-Related Problems and Staff Climate (SES measures); and in the student domain for Student Climate (SES and Student Surveys) and Physical Bullying Perpetration (TASB).

Inconsistencies among reporters existed, as well. For example, students in intervention schools perceived greater likelihood of staff intervention in bullying incidents; however, this effect was not found in the staff self-report (SES) data. This is not surprising, given that school staff tend to perceive themselves as being highly instrumental in bullying situations (Charach, Pepler, & Ziegler, 1995), possibly resulting in ceiling effects. Student perceptions of school climate were not significantly affected by program involvement, although school staff endorsed improvements in staff and student climate, and reductions in bullying perpetration were found only with teacher report, and only with regard to physical bullying. However, measures of bullying perpetration contained different items between reporters and we did not discriminate between physical (e.g., "I pushed, shoved, tripped, or picked a fight with a student who I knew was weaker than me") and nonphysical bullying (e.g., "I spread rumors about some students") with students as we did with teachers. It is not surprising that significant differences were not found between intervention and control schools on nonphysical bullying, because teachers have more difficulty discriminating between forms of bullying (Card, Stucky, Sawalani, & Little, 2008).

Results of this study stand in contrast to the paucity of evidence for effective school bullying preventive interventions in the United States. As recent reviews and meta-analyses have pointed out, the majority of efficacious school bullying prevention programs have been developed and tested in European countries (Baldry & Farrington, 2007; Ferguson, San Miguel, Kilburn, & Sanchez, 2007; Merrell et al., 2008; Smith et al., 2004; Ttofi, Farrington, & Baldry, 2008). Large-scale tests of United States school bullying prevention programs are few and have resulted in a weak

evidence base. This study relied on an explicated theory of change, well-developed intervention curricula, and psychometrically sound outcome measures to demonstrate the potential for efficacious school bullying prevention. Moreover, this study serves as a replication trial of the STR program, which was originally evaluated by Frey et al. (2005, 2009) in six Pacific Northwest elementary schools. In addition to using a larger number of elementary schools, the schools in this study were more diverse in terms of racial/ethnic makeup and geography. Our results are supportive of Frey et al.'s previous research (e.g., both studies found some positive effects on bystander behavior) and extend the generalizability of the program's efficacy to a broader range of schools.

Despite the emphasis we placed on designing and executing a rigorous school-randomized controlled trial, several limitations of this study should be mentioned. Unlike Frey et al. (2005), we did not use observational measures of playground behaviors despite the fact that their study found intervention effects in several bullying-related behaviors with this method. The use of students' self-report of bullying and bullying-related behaviors has been brought into question (Frey et al., 2009; Jenson & Dieterich, 2007). Although other sources of data (e.g., firsthand playground observations and direct observation of teachers in classrooms) could provide a unique perspective to program implementation and effectiveness, these are expensive and difficult to implement in large-scale school-randomized trials. Our study also was limited by its relatively short duration. As the STR program is designed to be administered to elementary students across three consecutive grades, long-term follow-up of students would allow for a better assessment of program sustainability. In addition, our assessment of program fidelity, as implemented in the classroom, was measured using online checklist data reported by teachers. Although analyses of these data indicated fidelity to the classroom curricula, observation of teachers administering the program in classrooms by trained observers could have provided an alternative assessment of classroom program fidelity, independent from teachers' self-report. We note that effect sizes presented in this study are relatively small and that the generalization of study findings also may be limited by selection of schools with at least 80% agreement from teachers to participate in the trial and the unique geographical and demographic features of the school sample recruitment area (i.e., northern-central California).

Effect for School Psychologists

Given an increasing societal concern regarding the prevalence and consequences of bullying in schools, school psychologists should be knowledgeable about the issue. It is often the school psychologist's charge to find comprehensive and effective programs to prevent student aggression and violence, and improve school climate and student learning. Results of this study facilitate a proactive approach to dealing with school bullying—an approach that allows school psychologists, teachers, and other school staff to work together in a coordinated fashion, across multiple levels of the school environment, to address school bullying. School psychologists, therefore, need to be aware of the evidence for effective school-based bullying prevention to guide these efforts, especially in light of the lack of strong evidence in the United States for school bullying prevention program effectiveness. Our findings support and extend previous research demonstrating evidence for the efficacy of the STR school bullying prevention program. Although school bullying-related behaviors by students are known to increase during the course of the school year, as was seen in this study, successful prevention program implementation can reduce this escalation. More research is needed to assess the long-term effects of the STR program, with full (3-year) program implementation, on distal mental health and academic student outcomes. Nonetheless, results from this study suggest that schools can take proactive steps even in the short run to address bullying in schools and provide school psychologists with

an important tool for combating school bullying.

References

Atlas, R. S., & Pepler, D. J. (1998). Observations of bullying in the classroom. *The Journal of Educational Research, 92,* 86–99.

Baldry, A. C., & Farrington, D. P. (2007). Effectiveness of programs to prevent school bullying *Victims & Offenders, 2,* 183–204.

Bloom, H. S., Bos, J. M., & Lee, S.-W. (1999). Using cluster random assignment to measure program impacts: Statistical implications for the evaluation of education programs. *Evaluation Review, 23,* 445–489.

Card, N. A., & Hodges, E. V. E. (2008). Peer victimization among schoolchildren: Correlations, causes, consequences, and considerations in assessment and intervention. *School Psychology Quarterly, 23,* 451–461.

Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, Intercorrelations, and relations to maladjustment. *Child Development, 79,* 1185–1229.

Charach, A., Pepler, D. J., & Ziegler, S. (1995). Bullying at school: A Canadian perspective. *Education Canada* (Spring), 12–18.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Committee for Children. (2001). *Steps to Respect: A Bullying Prevention Program.* Seattle, WA: Author.

Cook, C. R., Williams, K. R., Guerra, N. G., Kim, T. E., & Sadek, S. (2010). Predictors of bullying and victimization in childhood and adolescence: A meta-analytic investigation. *School Psychology Quarterly, 25,* 65–83.

Csuti, N. (2008a). The Colorado Trust Bullying Prevention Initiative Student Survey. Retrieved from www.thecoloradotrust.org/repository/initiatives/pdfs/BPI/Evaluations/BPI%20Student%20Survey.pdf

Csuti, N. (2008b). The Colorado Trust Bullying Prevention Initiative Staff Survey. Retrieved from http://www.thecoloradotrust.org/repository/initiatives/pdfs/BPI/Evaluations/BPI%20Staff%20Survey.pdf

Espelage, D. L., & Horne, A. M. (2008). School violence and bullying prevention: From research-based explanations to empirically based solutions. In S. D. Brown & R. W. Lent (Eds.), *Handbook of counseling psychology* (4th ed., pp. 588–606). Hoboken, NJ: John Wiley & Sons.

Espelage, D. L., Mebane, S. E., Swearer, S. M., Espelage, D. L., & Swearer, S. M. (2004). Gender differences in bullying: Moving beyond mean level differences. In D. L. Espelage & S. M. Swearer (Eds.), *Bullying in American schools: A social-ecological perspective on prevention and intervention.* (pp. 15–35). Mahwah, NJ: Lawrence Erlbaum.

Farrington, D. P., & Ttofi, M. M. (2009). *Campbell Systematic Reviews: School-based programs to reduce bullying and victimization.* Retrieved from http://www.campbellcollaboration.org/lib/download/718/

Ferguson, C. J., San Miguel, C., Kilburn, J. C., Jr., & Sanchez, P. (2007). The effectiveness of school-based anti-bullying programs: A meta-analytic review. *Criminal Justice Review, 32,* 401–414.

Frey, K. S., Hirschstein, M. K., Edstrom, L. V., & Snell, J. L. (2009). Observed reductions in school bullying, nonbullying aggression, and destructive bystander behavior: A longitudinal evaluation. *Journal of Educational Psychology, 101,* 466–481.

Frey, K. S., Hirschstein, M. K., Snell, J. L., Edstrom, L.V.S., MacKenzie, E. P., & Broderick, C. J. (2005). Reducing playground bullying and supporting beliefs: An experimental trial of the Steps to Respect Program. *Developmental Psychology, 41,* 479–491.

Glew, G. M., Fan, M. Y., Katon, W., Rivara, F. P., & Kernic, M. A. (2005). Bullying, psychosocial adjustment, and academic performance in elementary school. *Archives of Pediatrics & Adolescent Medicine*, *159*(11), 1026–1031.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60,* 549–576.

Graham, S., & Jaconet, J. (2002). Ethnicity, peer harassment, and adjustment in middle school: An exploratory study. *Journal of Early Adolescence, 22,* 173–199.

Hawker, D. S. J., & Boulton, M. J. (2000). Twenty years' research on peer victimization and psychosocial maladjustment: A meta-analytic review of cross-sectional studies. *Journal of Child Psychology and Psychiatry, 41,* 441–455.

Haynie, D. L., Nansel, T., Eifel, P., Crump, A., Saylor, K., Yu, K., et al. (2001). Bullies, victims, and bully/victims: Distinct groups of at-risk youth. *The Journal of Early Adolescence, 21,* 29–49.

Jenson, J. M., & Dieterich, W. A. (2007). Effects of a skills-based prevention program on bullying and bully victimization among elementary school children. *Prevention Science, 8,* 285–296.

Kaltiala-Heino, R., Rimpelä, M., Rantanen, P., & Laippala, P. (2001). Adolescent depression: The role of discontinuities in life course and social support. *Journal of Affective Disorders, 64,* 2–3.

Kochenderfer, B. J., & Ladd, G. W. (1996). Peer victimization: Cause or consequence of school maladjustment? *Child Development, 67,* 1305–1317.

Livert, D., Rindskopf, D., Saxe, L., & Stirratt, M. (2001). Using multilevel modeling in the evaluation of community-based treatment programs. *Multivariate Behavioral Research, 36,* 155–183.

Low, S., Frey, K., & Brockman, C. (2010). Gossip on the playground: Changes associated with universal intervention, retaliation beliefs, and supportive friends. *School Psychology Review, 39,* 536–551.

Merrell, K. W., Gueldner, B. A., Ross, S. W., & Isava, D. M. (2008). How effective are school bullying intervention programs? A meta-analysis of intervention research. *School Psychology Quarterly, 23,* 26–42.

Nansel, T. R., Haynie, D. L., & Simons-Morton, B. G. (2003). The association of bullying and victimization with middle school adjustment. In M. J. Elias & J. E. Zings (Eds.), *Bullying, peer harassment, and victimization in the schools: The next generation of prevention* (pp. 45–61). Philadelphia: The Haworth Press.

Nansel, T. R., Over peck, M., Pillar, R. S., Roan, W. J., Simons-Morton, B. G., & Scheldt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. *Journal of the American Medical Association, 285,* 2094–2100.

National Center for Education Statistics. (n.d.) [Data file]. Accessed June 18, 2008, from http://www.ed-data.k12.ca.us

Olweus, D. (1992). Bullying among schoolchildren: Intervention and prevention. In R. V. Peters, R. McMahon, & V. L. Quinsey (Eds.), *Aggression and violence throughout the life span* (pp. 100–125). Thousand Oaks, CA: Sage.

Olweus, D. (1993). Bully/victim problems among schoolchildren: Long-term consequences and an effective intervention program. In S. Hodgins (Ed.), *Mental disorder and crime* (pp. 317–349). Thousand Oaks, CA: Sage.

Porter, L., Batsche, G., Castillo, J., & Witte, R. (2006, March). *Problem-solving and response-to-intervention: School psychologists' beliefs, practices, and training needs.* Paper presented at the National Association of School Psychologists Annual Convention, Anaheim: CA.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. T., Jr. (2004). *HLM 6, hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International.

Raudenbush, S. W., & Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics, 20,* 307–335.

Rigby, K. (2001). Health consequences of bullying and its prevention in schools. In J. Juvonen & S. Graham (Eds.), *Peer harassment in school: The plight of the vulnerable and victimized* (pp. 310–331). New York: Guilford Press.

Ryan, W., & Smith, J. D. (2009). Antibullying programs in schools: How effective are evaluation practices? *Prevention Science, 10,* 248–259.

Schafer, J. L. (1997). *NORM: Multiple imputation of incomplete multivariate data under a normal model. Windows version 2.03.* University Park, PA: The Methodology Center at Penn State.

Seals, D., & Young, J. (2003). Bullying and victimization: Prevalence and relationship to gender, grade level, ethnicity, self-esteem, and depression. *Adolescence, 38,* 735–747.

Smith, J. D., Schneider, B. H., Smith, P. K., & Ananiadou, K. (2004). The effectiveness of whole-school antibullying programs: A synthesis of evaluation research. *School Psychology Review, 33,* 547–560.

Swearer, S. M., Espelage, D. L., Vaillancourt, T., & Hymel, S. (2010). What can be done about school bullying? Linking research to educational practice. *Educational Researcher, 39,* 38–47.

Ttofi, M. M., Farrington, D. P., & Baldry, A. C. (2008). *Effectiveness of programmes to reduce school bullying. Report prepared for the Swedish National Council for Crime Prevention.* Stockholm.

Vreeman, R. C., & Carroll, A. E. (2007). A systematic review of school-based interventions to prevent bullying. *Archives of Pediatrics and Adolescent Medicine, 161,* 78–88.

Eric C. Brown is an assistant research professor with the Social Development Research Group, School of Social Work, University of Washington. His current research includes applying advanced methods and designs to investigate the efficacy of school- and community-based preventive interventions. He is the principal investigator of the school-randomized controlled trial of Steps to Respect: A Bullying Prevention Program.

Sabina Low is an assistant professor of clinical psychology at Wichita State University. She was a co-investigator on a randomized controlled trial of Steps to Respect: A Bullying Prevention Program and is currently a principal investigator on a randomized controlled trial of a middle school bullying prevention program. She has published and presented on topics including adolescent psychopathology, bullying, and adolescent health risk behaviors.

Brian H. Smith completed his doctoral studies at the Social Development Research Group, School of Social Work, University of Washington, in 2004. Since that time he has been a research scientist at the Committee for Children, where his work is focused on the development and evaluation of school-based prevention and youth development programs.

Kevin P. Haggerty is the assistant director of the Social Development Research Group, School of Social Work, University of Washington. He specializes in the development and implementation of prevention programs at community, school, and family levels. He is an investigator on the Raising Healthy Children project and the principal investigator of the Family Connections study.

COPYRIGHT INFORMATION